The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed P. Mahoney

Dated 16 June 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
The Patent Office
2 JUL 2003
LONDON

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

1. **Your reference** 9.14.81083

03JUL03 E819735-1 D00027
P01/7700 0.00-0315525.6

02 JUL 2003

2. **Patent application number** *(The Patent Office will fill in this part)* 0315525.6

3. **Full name, address and postcode of the or of each applicant** *(underline all surnames)*

Pär Nordlund
Roslagsgatan 47
113 54 Stockholm
Sweden 08666125001

Tobias Cornvik
Fyrskeppsvägen 118
121 54 Johaneshov
Sweden 08666133001

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give country/state of incorporation

4. **Title of the invention** A Method of Screening

5. **Name of your agent** *(if you have one)* Frank B. Dehn & Co.

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

179 Queen Victoria Street
London
EC4V 4EL

Patents ADP number *(if you know it)* 166001

6. **If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and** *(if you know it)* **the or each application number**

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. **If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application**

Number of earlier application	Date of filing *(day / month / year)*

8. **Is a statement of inventorship and of right to grant of a patent required in support of this request?** *(Answer 'Yes' if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Yes

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	24
Claim(s)	2
Abstract	-
Drawing(s)	14 + 14

10. If you are also filing any of the following, state how many against each item.

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. I/We request the grant of a patent on the basis of this application.

Signature Frank B Dehn & Co Date 2 July 2003

12. Name and daytime telephone number of person to contact in the United Kingdom

Rebecca Gardner
020 7206 0600

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*
b) *Write your answers in capital letters using black ink or you may type them.*
c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s) of the form. Any continuation sheet should be attached to this form.*
d) *If you have answered 'Yes', Patents Form 7/77 will need to be filed.*
e) *Once you have filled in the form you must remember to sign and date it.*
f) *For details of the fee and ways to pay please contact the Patent Office.*

81083.621

A method of screening

The invention relates to a method of screening protein molecules, in particular to methods for identifying soluble proteins.

The production of pure or semi-pure proteins is important in many commercial and academic research and development programs. Often such proteins are produced recombinantly. Recombinant proteins can constitute products (e.g. enzymes for use in biomedical assays or in industrial processes) and are also used in the process of developing pharmaceutical drugs. In pharmaceutical drug development processes, these proteins are often used for structural studies (where methods such as NMR and x-ray crystallography are employed), and in biochemical or biophysical studies of the target protein. In academic research, recombinantly expressed proteins are used for biochemical, biophysical and structural characterisation. Additionally, the sequencing of the human genome has now revealed many potential targets for recombinant expression to elucidate the function of gene products.

Traditionally, recombinant proteins are produced by the overexpression of a gene of interest. However, many proteins aggregate (e.g. in inclusion bodies) when overexpressed and fail to fold into their native conformation. Such protein aggregates must be dissolved and correctly refolded before they can be used in many of the above methods. However refolding proteins from inclusion bodies usually results in very low yields of proteins and often it is not possible to determine if the protein isolated is in fact correctly folded.

To improve the yields of proteins obtained, it is therefore important for proteins to have a high solubility. Several studies have now shown that protein solubility may be substantially increased by amino acid

substitutions at particular positions. Mutant libraries of target proteins have been created in order to select those which have increased solubility and hence are capable of being highly expressed. Indeed, proteins which are naturally insoluble, may be mutated and expressed in soluble form. The process of mutating proteins for isolating/detecting soluble variants is known in the art as directed evolution. Several methods of mutagenesis are available including site directed mutagenesis, truncation of the sequence ends, use of an exonuclease enzyme and introduction of a randomised cassette of nucleotides into the nucleic acid sequence.

Mutant libraries of proteins must further be screened to detect the recombinant clone(s) containing soluble variants. Several screening methods have been described, which select proteins with increased solubility. Maxwell et al. (Protein Science, 1999, 8, 1908-1911), described a simple assay for assessing solubility using chloramphenicol acetyltransferase (CAT) fusion proteins, this was based on the principle that cells expressing fusions of an insoluble protein to CAT exhibit decreased resistance to chloroamphenicol compared to fusions with soluble proteins. Experiments were carried out using the wild-type (insoluble) catalytical core domain of HIV integrase and a soluble variant with amino acid substitutions at positions F185. Selection experiments on a library of proteins were not carried out using this screen, and the authors hypothesise that a large number of false positives could arise. Hence, it is clear that more reliable screens for soluble proteins are required.

Another group, Waldo et al. (Nature Biotechnology, 1999, 17, 691-695), have developed a soluble protein screen using proteins with N-terminal fusions to green fluorescent protein (GFP). Waldo et al. demonstrated that the correct folding of the GFP protein domain (and hence its ability to fluoresce) is directly related to

the folding robustness and avoidance of inclusion body formation of the protein of interest.

The correlation reported between non-fusion solubility and GFP fusion fluorescence has however been reported as not perfect, where solubility can be both over and under estimated. Generally, it will be recognised that techniques which rely on the properties of a protein fused to a reporter moiety will not always give a good indication of how the target protein free of its fusion partner will behave.

Peabody and Al-Bitar (Nucleic Acids Research, 2001, 29, No. 22 e113 1-7) developed a soluble protein screen, wherein recombinant bacterial colonies were overlayed with agarose gel. Proteins diffused from the colonies through the overlayed gel depending on molecular weight. Soluble variants which diffused faster gave larger diameter spots of greater intensity than insoluble variants. Diffusion takes 24 hours and so this method is not suitable for high throughput screening..

In Knaust and Nordlund (Analytical Biochemistry, 2001, 297, 79-85), a screen for soluble proteins was developed using filtration of cells lysed in culture in 96 well plates. Recombinant bacteria were grown on agar plates and then used to inoculate media in 96 well plates. The cultures were grown, lysed and filtered. Soluble proteins pass through a filter into a microtiter plate, positive recombinant bacteria were identified by screening the filtrates by Western blot.

Although this method proved to be reliable for soluble protein detection, it was also time consuming to process large numbers of recombinant clones due to the many pipetting steps necessary.

The present inventors have now developed a screening method which can process large numbers of recombinant bacteria in a short period of time. Surprisingly, it has been found that lysis of cells and filtration of lysates can be carried out directly on

colonies of cells, hence eliminating the requirement to grow colonies in culture and the multiple pipetting steps involved with this technique.

The present invention is hence capable of operating on large numbers of variants, is inexpensive and has a high reliability of predicting soluble variants.

Thus in one aspect the present invention provides a method of treating one or more colonies of cells which method comprises:

(a) subjecting the cell colonies to conditions which are capable of causing lysis thereof; and

(b) filtering the lysate of step (a) through a filter having pores which allow only soluble proteins to pass through the filter.

Preferably this method is used as a screening method to identify colonies expressing soluble proteins.

As used herein, the term "colony" or "colonies" describes a circumscribed group of cells, normally derived from a single cell or small cluster of cells, growing on a solid or semi-solid medium. Colonies can be formed from any cell type which can be made to express recombinant proteins and which can grow on solid or semi solid media. For example, colonies can be formed of prokaryotic e.g. bacteria or eukaryotic cells e.g. yeast, unicellular eukaryotes such as Leishmania, insect cells or mammalian cells or cell lines. Preferably colonies are formed of *E. coli, Bacillus subtilis, Streptococcus lactis, Streptomyces lividens, Lactococcus lactis, Staphylococcus aureas, Aspergillus niger, Picia pastoris, Saccharomyces cerevisiae or Schizosaccaromyces pombe*.

Semi-solid or solid media used to grow colonies typically consists of culture media with the addition of 0.1% or greater agar. More preferably, solid or semi-solid media contains at least 0.2%, e.g. at least 1.0% or at least 1.5% agar. "Colonies" do not encompass cells grown in liquid culture.

The lysis step of the present invention can be carried out chemically or otherwise using reagents which are well known in the art e.g. urea, lysozyme containing buffers or detergents. The degree of lysis must be sufficient to allow the proteins of the cell to pass freely out of the cell. Typically, when dealing with membrane bound proteins, lysis is performed in the presence of detergents or amphiphiles, for example Triton X-100 or dodecylmaltoside, to release the protein from the membrane. Preferably, the lysis step is non-denaturing, allowing proteins to retain a native, i.e. correctly folded or native-like conformation, this is referred to herein as 'native lysis'. The lysis step can alternatively be carried out by freeze thawing the colonies. More preferably, lysis is carried out using both native lysis buffer and freeze thawing the colonies. Preferably, the native lysis buffer contains lysozyme, for example at 50-750 μg/ml, more preferably at 100-200 μg/ml. DNAse can also be found in native lysis buffer preferably at 250-750 μg/ml. Native lysis buffer may contain, for example 20 mM Tris, pH 8, 100 mM NaCl, lysozyme (200 μg/ml) and DNAse I (750 μg/ml).

Typically, the colonies will be exposed to the lysis mixture (buffer) for 15-60 minutes, preferably around 30 minutes. The step of freeze thawing is preferably repeated, i.e. two or more cycles, preferably 3 or more cycles of freeze thawing are performed. In one preferred embodiment lysis is achieved by a 30 minute incubation at room temperature with lysis buffer and three x 10 minutes freeze thawing.

Typically, the percentage of cells lysed within a colony during the lysis step is 5-50%. Colonies to be lysed preferably contain at least 10^4 cells e.g. at least 10^5, at least 10^6, at least 10^7 or at least 10^8 cells. The size of such colonies is typically 0.1-3 mm^2, preferably 0.2-2 mm^2 and more preferably 0.25-1 mm^2. However, it will be appreciated that the present method

could be used to screen a wide range of colony sizes.

A "native-like" protein or "native" protein refers to a soluble intracellular, extracellular or membrane protein wherein the protein exhibits a native-like conformation and functions similarly or identically to the naturally occurring protein. "Native" or "native-like" proteins are expressed in soluble form and/or are correctly folded. Native-like membrane proteins do not have to be present free in solution, but may be present in membrane vesicles rather than inclusion bodies. Thus "native-like" proteins are generally not insoluble, present in inclusion bodies, aggregated or misfolded. According to the methods of the present invention, it is the soluble proteins which are able to pass through the filter and thus be separated from insoluble proteins. As described herein, this allows the identification of colonies expressing soluble proteins, e.g. through the use of blotting techniques. The correlation between solubility and correct, i.e. native or native-like folding means that the method is able to separate proteins with their native conformation from misfolded/aggregated proteins.

A "soluble" protein can thus be defined with reference to possession of a native or native-like conformation. Further, a soluble protein can be described as a protein which remains in the supernatant after cell lysis and centrifugation thereof. Centrifugation can be carried out at at least 1000g, preferably at at least 3000g, preferably at at least 10000g and more preferably at around 20000g. Centrifugation can be carried out at 100000g. The duration of centrifugation can be from 1 minute (typically at least 10 mins) to at least 1 hour, where the duration required generally decreases as the centrifugal force increases. Particularly suitable conditions for providing only soluble proteins in the resultant supernatant include 10 minutes at 100000g, 30

minutes at 3000g or 15 minutes at 20000g. 15 minutes at 20000g as described in the Examples being especially suitable.

Some proteins which pass through the filter will not correspond exactly to any naturally occurring protein. For example a library of proteins may be generated based on a target protein in order to identify related mutants with improved solubility in a given expression system compared to a problematic target protein. In these circumstances it is appropriate to consider the soluble mutants as having a "native-like" conformation.

The filtration step of the invention can be carried out using standard filter membranes for the filtering of biological samples. The filters will typically have a pore size from 0.015 μm to 12 μm, preferably from 0.35 to 1.2 μm, preferably from 0.45 to 1.2 μm, more preferably from 0.45 to 0.8 μm. Preferably, the filters have pore sizes below 4.0 μm, typically below 2.0 μm, more preferably below 1.0 μm. For several cell types, in particular bacteria e.g. *E. coli*, an optimal pore size may be 0.1-1.5 μm. For eukaryotic cells, preferred pore sizes may be larger. It will be appreciated that filters are manufactured and marketed as having a particular pore size but the manufacturing process may occasionally result in a few smaller or larger pores; the sizes listed, which refer to the diameter, are thus the most common pore size of a given filter. Although reference is made to a range of potential pore sizes, any single filter will usually have one designated pore size, e.g. 0.45 μm. Suitable filters are Super and GH polypro (from Pall) and Nucleopore (from Whatman).

It will be appreciated that different cell types may require the use of filters with different pore sizes, due for example to their different tendencies to harbour aggregated proteins, which aggregates may also have varying properties in different cell types.

Selection of a suitable filter is well within the competency of someone skilled in this field. For example, it is possible to select an appropriate pore size, by using a set of test proteins for the desired cell type and investigating their behaviour with filters of varying pore sizes. Such data should then be compared with centrifugation data as shown in Example 1.

Preferably the filter is overlayed on the colonies to lift the colonies/protein from the semi-solid or solid growth media (Figure 1). Alternatively, filters could be placed on the growth media and cells seeded directly onto the filter, the filter could then simply be lifted off with the colonies already on it. Preferably, the lifting of colonies from their growth media can be carried out prior to the lysis step. The lysis can hence be carried out directly on colonies on a filter. The filter with colonies attached can be treated with lysis buffer (Figure 2) or overlayed on other membranes/ filters treated with lysis buffer and/or subjected to freeze/thaw treatments.

Filtration is carried out after lysis, i.e. it is the lysate which is filtered. It will be appreciated however that filtration and lysis may occur simultaneously when considering a whole colony since some cells may undergo lysis before others and hence may be filtered before or at the same time as others are lysed. Preferably, proteins which pass through the filter are held on a capture membrane, to allow screening/detection of any proteins of interest and then to allow the identification of colonies expressing such proteins. Such capture membranes may typically comprise nitrocellulose. In a preferred embodiment, proteins can simply be allowed to pass through the filter, possibly as a result of an active capillary action. In another embodiment, force may be applied to aid filtration. The force can be applied vertically on the filter paper, wherein such forces can include the application of

pressure or a vacuum.

The capture membrane can fix the soluble proteins from the individual colonies and their positions on the capture membrane can then be compared to the filter carrying the original colonies. Thus, from the colony filtration blot, it is possible to track back and identify the original colonies expressing the soluble proteins of interest on the growth media. To aid in the process of identifying clones expressing soluble proteins, positive controls can be used. These are clearly seen on the final colony filtration blots and can enable the membrane/blot to be correctly orientated with the original colonies (Figs. 7 and 8).

In another embodiment, the filter with colonies attached can be placed colony side down onto a material soaked in lysis buffer. A (nitrocellulose) capture membrane can then be placed on top of the filter with colonies and several layers of filter paper (and paper towels) can be placed on top of this (Figure 3a). Force can then be applied to the top of this "sandwich" and ideally transfer buffer poured around the bottom to facilitate filtration and transfer of proteins onto the capture membrane.

In another embodiment the filter is placed colony side up onto a capture membrane and a vacuum is applied to "pull" protein through the filter paper and onto the capture membrane (Figure 3b).

In practice, lysis and filtration may conveniently take place in one overall step, e.g. during the application of conditions capable of causing lysis (e.g. 3 periods of freeze thawing), the cell lysate is filtered and captured on a capture membrane.

Alternatively viewed, the present invention provides a method of separating soluble from insoluble proteins, which method comprises:

(a) subjecting one or more colonies of cells to conditions which are capable of causing lysis thereof;

(b) filtering the lysate of step (a) through a filter having pores which allow only soluble proteins to pass through the filter, thereby generating a filtrate containing soluble proteins. According to this method, most or all insoluble proteins will fail to pass through the filter and hence separation of soluble from insoluble proteins occurs.

The methods of the present invention can also include the detection of a protein of interest after filtration. Proteins of interest can be detected using various tags which are well known in the art, e.g. histidine tag, VS tag, T7 tag, FLAG tag or any short protein sequence to which a specific antibody is available, glutathione-S-transferase, thioredoxin, green fluorescent protein and maltose-binding protein. Tags can be attached to a protein of interest generally by expressing such proteins as fusion proteins. As such, short tags are preferred, to allow proteins of interest to maintain a native-like conformation. Further, C-terminal tags are preferred, although N-terminal His tags are also used. Proteins of interest can also be detected using antibodies, monoclonal or polyclonal, either directed to a tag or directly to the protein of interest (expressed on its own or as a fusion). Proteins can also be detected if an enzymatic activity is exhibited, for example fusion tags that possess enzymatic activity include green fluorescent protein, horseradish peroxidase and glutathione-S-transferase. Further, proteins may be selected if they themselves have enzymatic activity. If the different colonies are each representative of a particular variant from a generated library, proteins of interest will typically be soluble members of the library. Conveniently therefore the members of the library will be expressed as fusions with a small tag to aid detection or antibodies to the library members may be used.

The filtrates of soluble proteins can further be

used in assays to test for the biochemical activity of the protein of interest. In a preferred embodiment, the filtrate can be simultaneously screened for the amount of a soluble protein of interest (e.g. using a tag) and for the activity of that protein using a suitable assay.

Upon identification of a positive colony (i.e. one expressing a target protein in soluble form), such colonies can be cultured and soluble or total protein isolated for Western blotting if confirmation of the screening results are required. Such clones can be used for the overexpression of the protein of interest for many different purposes, e.g. for structural studies to elucidate the protein sequence.

The methods of the invention are of utility in the separation of soluble and insoluble proteins. Soluble proteins can often be expressed in greater amounts and can be used in techniques such as NMR/X-ray crystallography for structural genomics. The methods of the present invention can also be used to screen cDNA libraries for particular clones expressing soluble proteins. The invention can also be used to screen for soluble variants of a particular protein.

Filtration is performed so that the majority of proteins passing through the filter, in particular the majority of a target protein or proteins passing through the filter are in soluble form.

Genes/cDNAs/coding regions encoding a protein of interest can be mutated to produce variants of that protein with varying degrees of solubility. These mutants can be produced in an expression system, wherein the most soluble variants can be selected using the lysis and filtration steps of the present methods performed on transformed colonies. Genes/cDNAs/coding regions can be transformed or transfected into expression systems in vectors/constructs, such as plasmids, viral vectors, cosmids and YACs. Such vectors may contain regulatory sequences and other elements well

known in the art. For example, the gene/cDNA/coding region may be placed under the control of a promoter in a vector. Promoters used are generally capable of expressing the protein of interest within a particular host. In a specific embodiment, the promoter used is inducible i.e. the expression of the protein of interest can be controlled. Such inducible promoters/systems include lac wherein induction of expression is controlled by the addition of IPTG and tet on/off, wherein induction of expression is controlled by the presence/absence of tetracycline and others are known in the field.

Many different methods of mutagenesis are known in the art which could be employed to create a library of variants of a protein of interest. Possible procedures include truncation of the sequence, use of an exonuclease enzyme, introduction of a randomised cassette or site-directed mutagenesis. For truncation, the number of nucleotides removed may be less than 2000, preferably less than 1000, and more preferably less than 800. Introduction of a randomised cassette for mutagenesis preferably uses a cassette containing less than 100 nucleotides.

Mutagenesis is preferably carried out on several copies of a nucleic acid sequence encoding a protein of interest so that a set of different mutated sequences can be screened, hence increasing the probability of identifying a native-like protein with improved solubility. The use of random mutagenesis is especially preferred where there is no prior knowledge of which particular mutations may yield a soluble variant.

Libraries of proteins can be created where the coding region has been randomly mutagenised and where different length constructs have been generated by erase-a-base or random priming reactions.

Further, vectors can be randomly mutagenised (or vector libraries), preferably in the promoter region. N

or C terminal tags or in the origin of replication and hosts can be randomly mutagenised, for example by random knockouts or from preselected libraries of strains. The methods of the invention can hence be used to screen directly for the expression of soluble protein from a limited number of genes (transferring cells with just cloned genes) or to screen large numbers of genes e.g. from cDNA expression libraries.

In a further embodiment, the present invention provides a kit for use in a method of treating one or more colonies of cells which comprises:

(a) a filter having pore sizes which only allow soluble proteins to pass through the filter;

(b) a capture membrane; and optionally

(c) reagents for use in native lysis of the cell colonies. As discussed above, (b) is preferably a nitrocellulose membrane and (c) preferably includes one or more components of a native lysis buffer as described herein. Suitable filters and capture membranes are also discussed herein.

The invention will now be further described in the following non-limiting Examples in which:

Figure 1 shows a method of peeling/lifting colonies from the semi-solid/solid growth media using a filter;

Figure 2 depicts one procedure to obtain colony lysis, wherein the filter with colonies is placed colony side up onto nitrocellulose and the filter soaked in native lysis buffer; following lysis of the colonies, soluble proteins can pass through the filter onto the membrane;

Figure 3a shows one embodiment of the invention where the "sandwich" method is employed. A filter with colonies is placed colony side down on filters soaked in lysis buffer. Nitrocellulose is placed on top of this and filter paper and paper towels on top of this. A force is then applied;

Figure 3b shows another embodiment of the invention

where a filter with colonies is placed colony side up on a nitrocellulose membrane in a column with buffer. A vacuum is then applied to aid filtration;

Figure 4a shows a colony filtration blot for total protein for all 48 different constructs used to screen for soluble proteins of 24 *E. coli* proteins;

Figure 4b (i) shows a colony filtration blot of the 48 constructs (24 different *E. coli* proteins with either an N-terminal His or flag tag), where positive colonies (hence soluble) can be seen; (ii) shows a table describing the expression levels of each protein in the colony blot, where these levels were compared to the values of the centrifugation blot (71% were predicted in the right category and only 8% were either false positives or false negatives);

Figure 4c (i) shows a dot blot for the same 48 constructs, wherein soluble protein fractions have been isolated by centrifugation; (ii) shows a table describing the expression levels of each protein in the centrifugation blot.

Figure 5 shows results of a dot blot comparing the number of positive GST tagged clones obtained when denaturing or native lysis buffer is used;

Figure 6 shows a schematic presentation of the Erase-A-Base process; the plasmid is linearized by endonuclease digestion in two unique restriction sites introduced in the cloning primer (2); the enzymes are chosen in such a way that one leaves a 3' overhang protected from ExoIII digestion and one leaves 5' which is susceptible to the digestion; samples from the ExoIII digestion are removed at timed intervals (3) and added to tubes containing S1 Nuclease, which removes the remaining single-stranded tails (4); the plasmid is religated (5) and transformed into a *E. coli* cloning strain (6,7); the library is then recovered from the cloning strain and transformed into an expression strain and screened for soluble expression;

Figure 7a shows examples of four different colony filtration blots; two blots that are from the RORa library and two from the SOCS-2 library; the time points where the aliquots were removed are noted under the pictures; the later time points do not contain any or very few colonies judged as positives; this is due to the fact that constructs at these time points correspond to proteins which have their start inside a domain;

Figure 7b shows RORA 5-8 minutes in more detail, where the positive controls are located and some examples of colonies judged as positive/soluble;

Figure 8 shows positive controls being added to a plated library; positive controls will form positive colonies and can serve as reference points to help identify other positive colonies expressing soluble protein;

Figure 9 shows Western blot of soluble protein fractions of 24 different clones picked from a filtration colony screen of RORa library; the cells were grown in liquid culture and lysed using freeze thawing; the insoluble material was pelleted using centrifugation at 20000g for 15 minutes; + indicates that the clone was judged as soluble from the screen; - indicates that the clone was judged as insoluble from the screen;

Figure 10 shows Western blot of eight clones identified as soluble from the colony filtration screen of a SOCS-2 library; the cells were grown in liquid culture and lysed using freeze thawing and the insoluble material was pelleted using centrifugation at 20000g for 15 minutes; all eight clones show soluble protein expression;

Figure 11 shows a schematic presentation of the protein domain organisation of RORa and SOCS-2; the green circles show the start position of clones identified as soluble in the colony filtration screen of the libraries; interestingly the start positions of

these clones are all located in between domain borders.

Figure 12 shows colony filtration blot for an *E. coli* integral membrane protein predicted to contain 13 transmembrane segments (amino acid permease) in a) native lysis buffer (as in Example 1) b) denaturing lysis buffer (+8M Urea) c) + Triton X-100 (1%) d) dodecylmaltoside (5 mM).

Examples

Example 1 - screening for soluble variants

24 *E. coli* proteins in two different expressions vectors, N-terminal His- or Flag-tag, with known solubility characteristics were used to test the Colony Filtration blot procedure (CoFi blot). The method uses an antibody-reaction for detection of soluble protein, and is therefore universally applicable to any protein containing a suitable tag, or other fusion polypeptide moiety, against which antibodies can be generated.

Materials and methods

Materials: Chemicals were from ICN (Costa Mesa, CA, USA) or Sigma-Aldrich Sweden (Stockholm, Sweden), if not stated otherwise. Oligonucleotides for PCR were purchased from Invitrogen.

Cloning of recombinant *E. coli* clones: DNA-fragments coding for 24 different *E. coli* proteins were amplified from genomic *E. coli* DNA by PCR, using specific primers. PCR products were cloned using the Gateway system (Invitrogen) into two pET based expression vectors containing either a N-terminal His-tag or a N-terminal Flag-tag. All constructs contain a C-terminal his-tag used for detection.

Small-scale expression of test clones: For control expression tests, the plasmids were freshly transformed into *E. coli* strain Bl21 (DE3), single colonies were picked and grown overnight in LB medium (Difco, Detroit, MI, USA) at 37°C while shaking at 250 rpm. For test expression, 1 ml LB was inoculated 1:10 with overnight culture and grown to an OD_{600} of about 0.6. Cultures were induced by addition of IPTG to a final concentration of 1 mM and grown for another 4 h in a shaker running at 250 rpm at 37°C. The cells were harvested by centrifugation at 2 000 g.

Culturing on LB media plates: Freshly transformed

cells from the 48 different constructs (24 different proteins in two expression vectors) were arrayed on two LB plates containing appropriate antibiotics at 37 degrees. The colonies were transferred to a Durapore filter membrane with 0.45 μm pore size (Millipore, Bedford, MA, USA) by gently applying the filter membrane on top of the LB plate, thereby putting the filter membrane in contact with the colonies. Through this procedure most of the colonies are transferred to the surface of the filter membrane. The filter membranes were then transferred with the colony side up to a new LB plates containing IPTG resulting in induced expression. After 4 hours the filter membrane containing the colonies was subjected to lysis. The procedure described was done for two membranes in parallel, one to be used for detection of soluble proteins and one to be used to detect total protein, ie the sum of the aggregated protein and the soluble protein.

Lysis of colonies and transfer to detection membrane: The lysis for total protein was done by placing one of the filter membranes on top of a "lysis and detection sandwich" constituting of one Protran BA 45 nitrocellulose membrane (Schleicher & Schuell, Dassel, Germany) and a 3MM paper (Whatman) drenched in denaturing lysis buffer. (8M Urea, 20 mM Tris, pH 8, 100 mM NaCl). The colonies are subjected to this treatment for 1 hour at room temperature.

For lysis under non-denaturing conditions the second membrane was placed on a similar sandwich soaked in native lysis buffer containing 20 mM Tris, pH 8, 100 mM NaCl, lysozyme (200 μg/ml = 2.35 Units/ml) and DNAse I (750 μg/ml = 487.5 Units/ml). The "lysis and detection sandwich" is then frozen at -80°C for 10 minutes and thawed for 10 minutes at 37 degrees. This freeze/thaw procedure is repeated 3 times.

Solubility assay, dot-blots: For solubility assays,

cell pellets from small-scale expression cultures were resuspended in lysis buffer consisting of 20 mM Tris, pH 8, 100 mM NaCl, lysozyme (100 μg/ml = 2.35 Units/ml), DNAse I (750 μg/ml = 487.5 Units/ml) and freeze/thawed 4 times. The lysates were then centrifuged at 20 000 g for 15 minutes for separation of the soluble protein in the supernatant from insoluble protein (inclusion bodies) precipitating in the pellet.

Samples from the soluble fractions were dotted onto a nitrocellulose membrane (Schleicher & Schuell, Dassel, Germany).

Antibody-incubation and development of colony filtration blots and dot blots: Colony filtration blots and dot blots were treated similarly after transfer of protein to nitrocellulose. Membranes were blocked in 1% BSA in TBST (20 mM Tris, pH 7.5, 500 mM NaCl, 0.05% Tween-20) for 1 h or overnight. The blots were then washed three times for 10 min in TBST. The membranes were incubated for 1 hour with INDIA His Probe (Pierece) diluted 5 000 in TBST. Blots were then washed three times for 10 min in TBST. The membranes were developed using a chemiluminescence solution SuperSignal WEST DURA (Pierce) and the images were captured with a CCD camera (Biorad). The level of expression was quantified using the TotalLab software (Pharmacia biotech)

Results

Confirmation of soluble protein expression by colony filtration blot

Total protein expression was confirmed for the 48 different constructs (Figure 4a). When the colon filtration blot (Figure 4b) was compared to the dot blot of the soluble protein fractions after centrifugation analysis (Figure 4c) only 4 out of 48 constructs were false positives or false negatives. The values of the integrated intensity of the dots were obtained from the TotalLab software. Using the intensities, the dots were

categorised into 4 different categorise; No, Low, Medium or High level of soluble protein expression 72% were predicted in the right category. 20% was predicted in the neighbouring category and only 6% in most distant. (Table A)

Example 2 - Construction and screening of deletion libraries:

To obtain deletion clones, a procedure was employed which has been described before in standard molecular Biology Protocol manuals such as Ausubel *et al* (Short Protocols in Molecular Biology, 2nd Ed, 1992, ISBN 0-471-57735-9). Here, a commercially available kit from Promega, Madison, WI, USA (the Erase-a-base® system) was used to perform this reaction step.

To create the library the vector containing the target gene is cut open with restriction enzymes just in front of the gene. The restriction enzymes for this are chosen so they produce suitable 5'- and 3'-ends. The 3'-end is protected from Exonuclease III attack whereas the 5'-end is susceptible to it. Exonuclease III is added and small aliquots at certain time points are removed from the reaction. These aliquots contain constructs of varying length that are religated and transformed into *E. coli*.

Two different proteins have been subjected to our deletion method coupled to the colony filtration blot.

RORa was until recently an orphan nuclear receptor with no known ligand. In late 2002 the structure of RORa was solved and the ligand was shown to be cholesterol (Kallen JA et. al.). SOCS-2 is involved in negative regulation of cytokine signalling and to date at least three different modulating mechanisms have been demonstrated, one which involves the SOCS-box targeting bound proteins to proteasomal degradation.

Material and Methods

The cDNA clones were chosen from the Mammalian Gene Collection database and ordered from RZDP Deutsches Ressourscenzentrum für Genomforschung.

Cloning

The targets were amplified using Touch down PCR with a plasmid containing the cDNA as template. The PCR primers were designed in such a way that suitable restriction sites were located in front of the open reading frame. The PCR products were cloned into a pET based expression vector using conventional restriction/ ligation methods.

Creation of deletion libraries

The pET based vectors containing the coding sequence for the two proteins were linearized by endonucelase digestion in the two unique restriction sites introduced in the cloning primer. RORa was digested with SphI and HindIII and SOCS-2 with SphI and XhoI both digestions were made in NEB buffer 2 and at 37 degrees for 3 hours.

The digests were verified by agarose gel electrophoresis and purified (QIAquick PCR Purification Kit, QIAGEN). The Erase-a-Base Kit from Promega was used according to the manufacturers instructions with the following alterations. The Exonuclease III digestions were made at 27 degrees to achieve a rate at of approximately 70 bp/min. 24 aliquots were removed from the ExoIII digestion mix, every minute for RORa and for SOCS-2 every 30 seconds. The 24 samples were pooled into six and treated with S1 nuclease. The DNA was precipitated using Ethanol and Ammoniumacetate. The precipitated DNA was resuspended in 10mM Tris pH 8 and treated with Klenow polymerase to flush the ends. The plasmids were religated and transformed into a cloning strain, DH5a or Top10 plated on LB plates containing the appropriate antibiotics. The library was recovered from the cloning strain by making a plasmid preparation from the colonies on the plate.

Screening for soluble protein expression

The deletion libraries were transformed into an expression strain BL21(DE3) and plated onto plates containing the appropriate antibiotics. Positive controls, constructs expressing soluble proteins, were also added to the plate to *facilitate orientation*. The plates were incubated at 37 degrees overnight. The screening for soluble protein was performed as described in *Example 1* with the following exceptions. Expression was induced for 6 hours at 25 degrees instead of 37 and 4 hours. For RORa 12 positive clones and 12 negative clones and for SOCS-2 14 positive clones were taken for further analysis.

Solubility assay, SDS PAGE and Western blot transfer

The clones that were picked for further analysis were characterized as described in *Example 1*. With the exception that the cells were grown for 6 hours at 25 degrees after induction of expression.

Antibody-incubation and development of colony blots and Western blots

The colony blots and western blots were developed as described in *Example 1*.

Results

Libraries completely covering RORa and SOCS-2 were made and screened for soluble protein expression using our soluble colony blot method. All positive clones (soluble) were confirmed as soluble by conventional methods, eg. centrifugation, gel electrophoresis and western blots (Figures 9 and 10). The clones judged as negative in screen did not give any bands in the Western blot (Figure 9). Interestingly when the sizes were compared to the PFAM domain structure all soluble clones were located in between predicted domain borders. The last soluble construct in the RORa library appear very close to start of the cholesterol-binding domain seen in

the crystal structure (Figure 11).

Example 3 - Constructions and Screening of a Library of Membrane Proteins

Cloning of recombinant *E. coli* clones: DNA fragments coding for 42 different *E. coli* membrane proteins were amplified from genomic *E. coli* DNA by PCR, using specific primers. PCR products were cloned into Gateway vectors by homologous recombination. The plasmids were transformed into *E. coli* strain C41. Single colonies were picked and grown overnight in LB medium (Difco, Detroit, MI, USA) at +37°C while shaking at 200 rpm. Cell stocks were prepared by adding glycerol to cell culture aliquots to the final concentration of 25%, and stored at -80°C.

Culturing on LB media plates: Cells from frozen stocks were resuspended in LB, plated on LB media plates containing an appropriate antibiotic, and grown overnight at 37°C. The colonies were transferred to filter membranes of different type and origin (Millipore, Bedford, MA, USA; or Pall Life Sciences, Ann Arbor, MI, USA), or of different pore size (0.45 to 3 μm) by gently placing the filter membrane on top of the LB plate, thereby putting the membrane in contact with the colonies.

The filter membrane were then transferred with the colony side up to a new LB plate containing IPTG for induced expression at +37°C for 3 hours.

Lysis of colonies and transfer to detection membrane: Lysis under native and denaturing conditions was performed as described in example 1, for 60-90 min.. Lysis in presence of detergents was performed in the same way as the native lysis in Example 1, except Triton X-100 (1%) or dodecylmaltoside (5mM) was added to the native lysis buffer. Detection was done using an

antibody specific for a C-terminal His-tag.

Results

The membrane proteins from the colonies which were lysed in a native lysis buffer with detergent present where clearly visible in the blots, while proteins from colonies lysed without detergent present were not visible (See Figure 12). The addition of detergent solubilised membrane proteins allowing them to pass the filtration membrane and to be visible on the colony blot.

Claims

1. A method of treating one or more colonies of cells which method comprises:

(a) subjecting the colonies to conditions which are capable of causing lysis thereof; and

(b) filtering the lysate of step (a) through a filter having pores which allow only soluble proteins to pass through the filter.

2. The method of claim 1 wherein the lysis of step (a) is native lysis.

3. The method of claim 1 or 2 wherein said filter has a pore size between 0.1 and 1.5 μm.

4. The method of any one of claims 1 to 3 wherein said colonies are lifted from their growth media on the filter used in step (b).

5. The method of claim 4, wherein said colonies are lifted prior to the lysis of step (a).

6. The method of any one of claims 1 to 5, wherein filtration step (b) includes the application of a force to the filter carrying the colonies.

7. The method of any one of claims 1 to 6, which further comprises:

(c) capturing proteins in the filtrate from filtration step (b) on a capture membrane.

8. A method of separating soluble from insoluble proteins, which method comprises:

(a) subjecting one or more colonies of cells to conditions which are capable of causing lysis thereof;

(b) filtering the lysate of step (a) through a

filter having pores which allow only soluble proteins to pass through the filter, thereby generating a filtrate containing soluble proteins.

9. The method of claim 8, wherein soluble proteins in the filtrate are identified using antibodies and/or fusion tags.

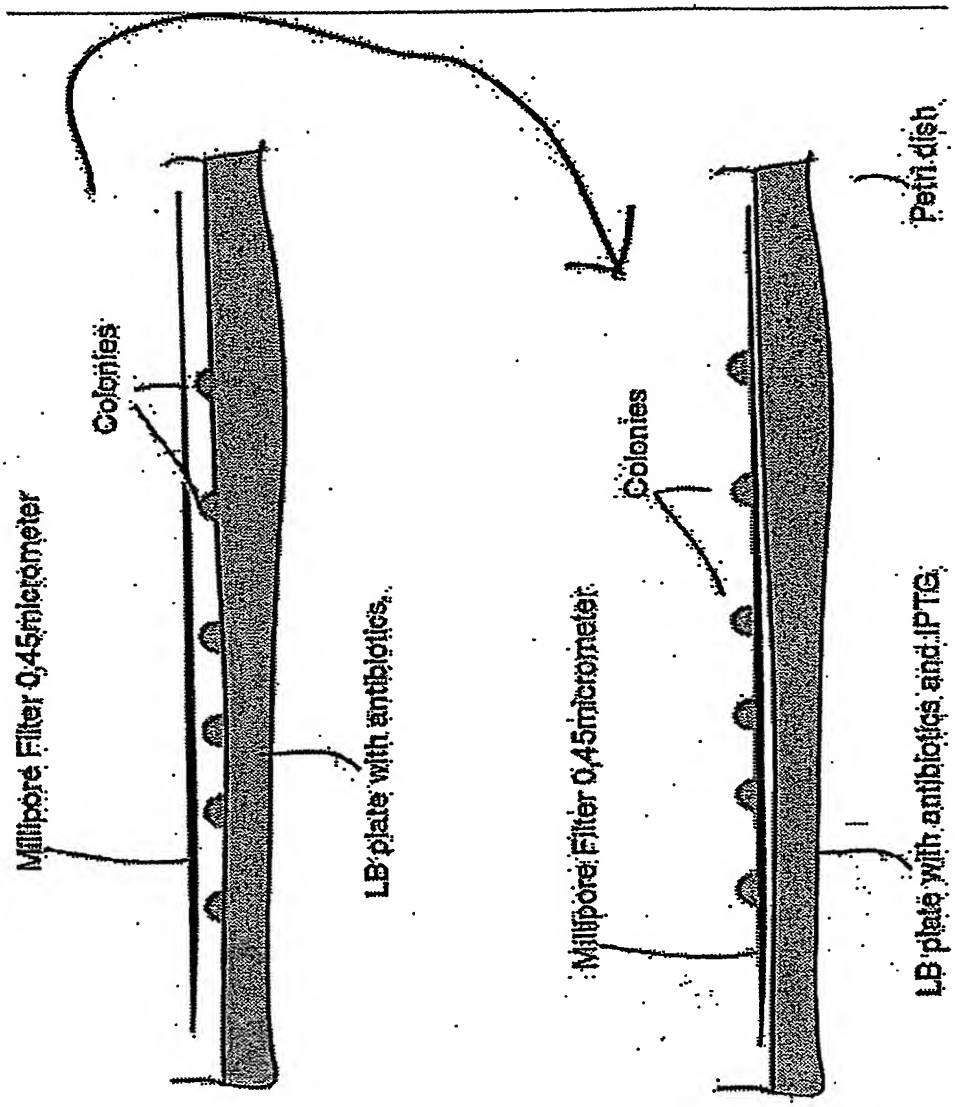

Millipore Filter 0,45micrometer
Colonies
LB plate with antibiotics
Millipore Filter 0,45micrometer
Colonies
Petri dish
LB plate with antibiotics and IPTG

FIGURE 1

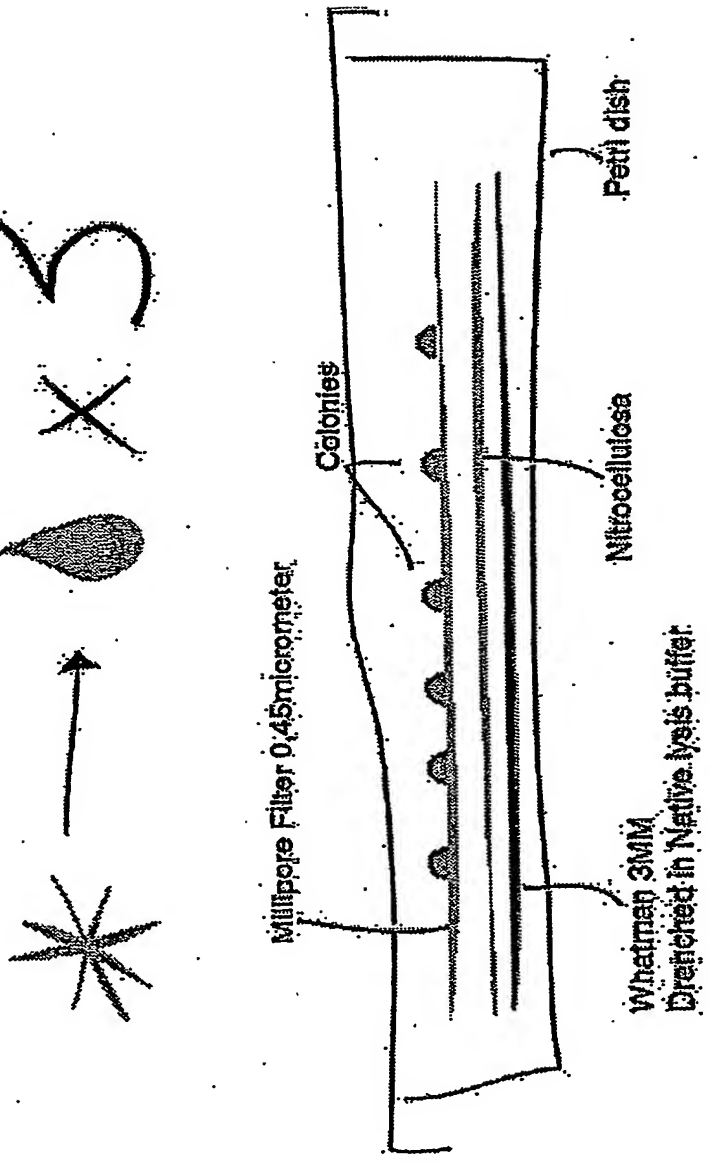
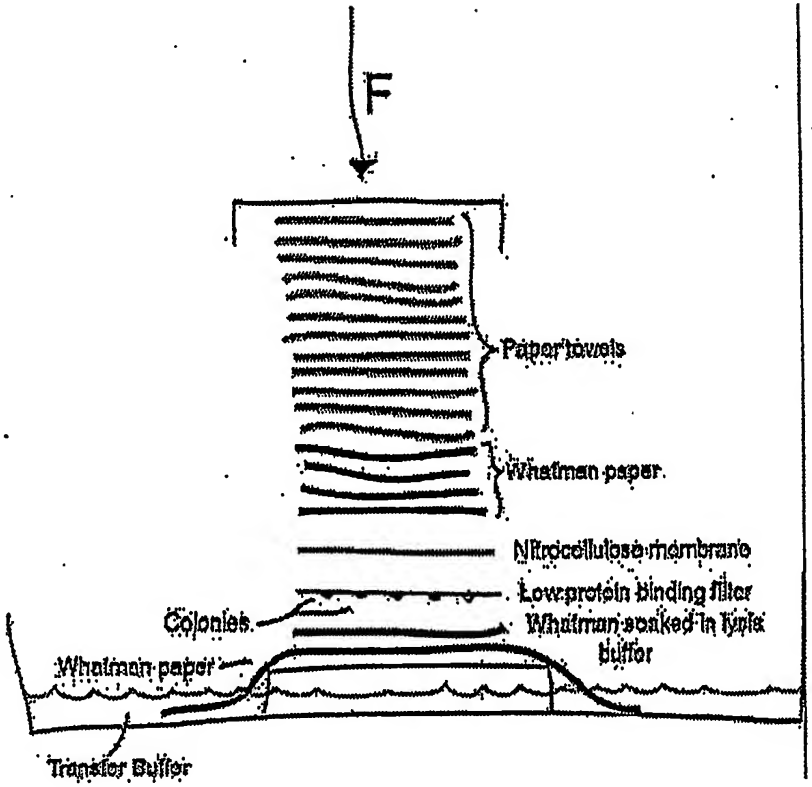
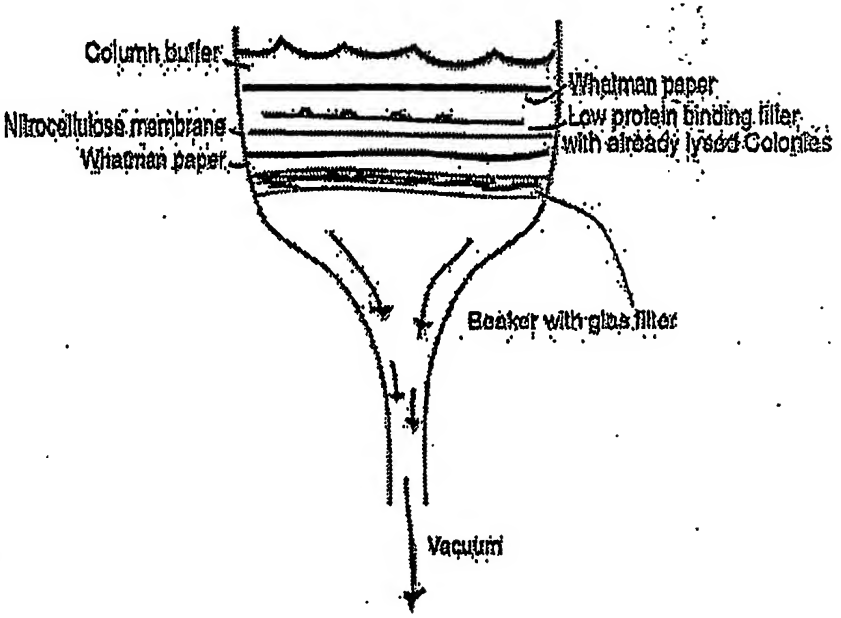

Millipore Filter 0,45micrometer
Colonies
Petri dish
Nitrocellulosa
Whatman 3MM
Drenched in Native lysis buffer

FIGURE 2



Sandwich "paper towel" method

FIGURE 3A



Vacuum

FIGURE 3B

Figure 4a

Total protein fraction (8M Urea)

His-tag

Flag-tag

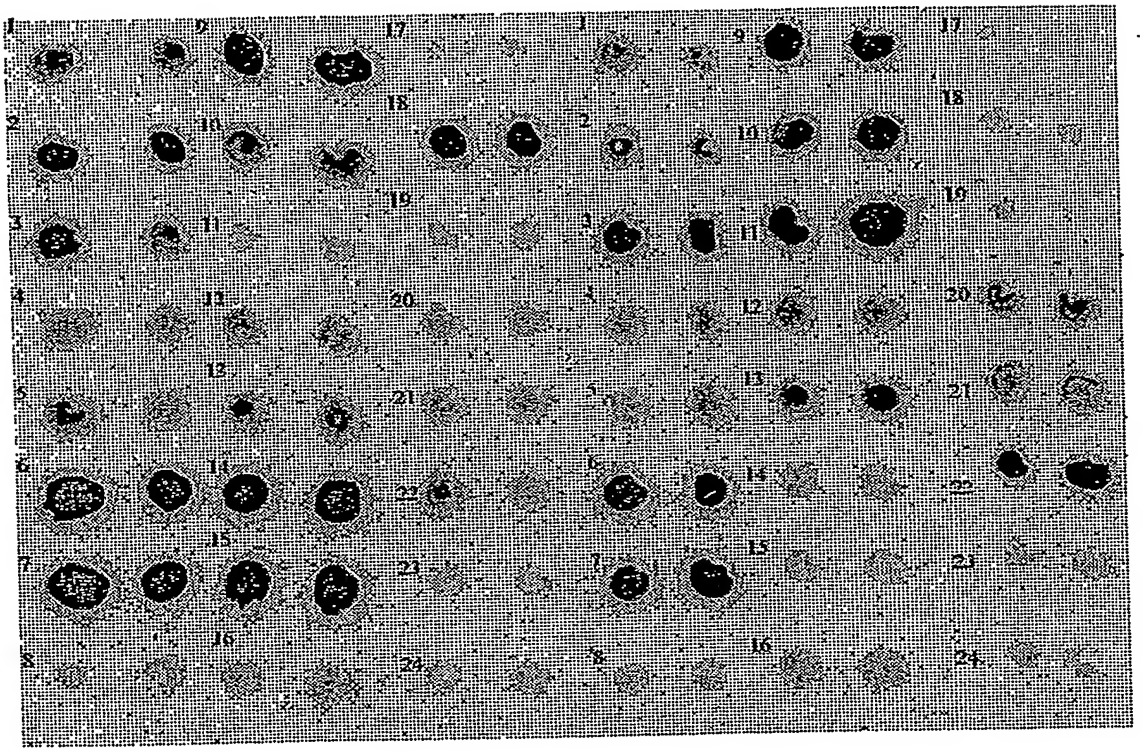
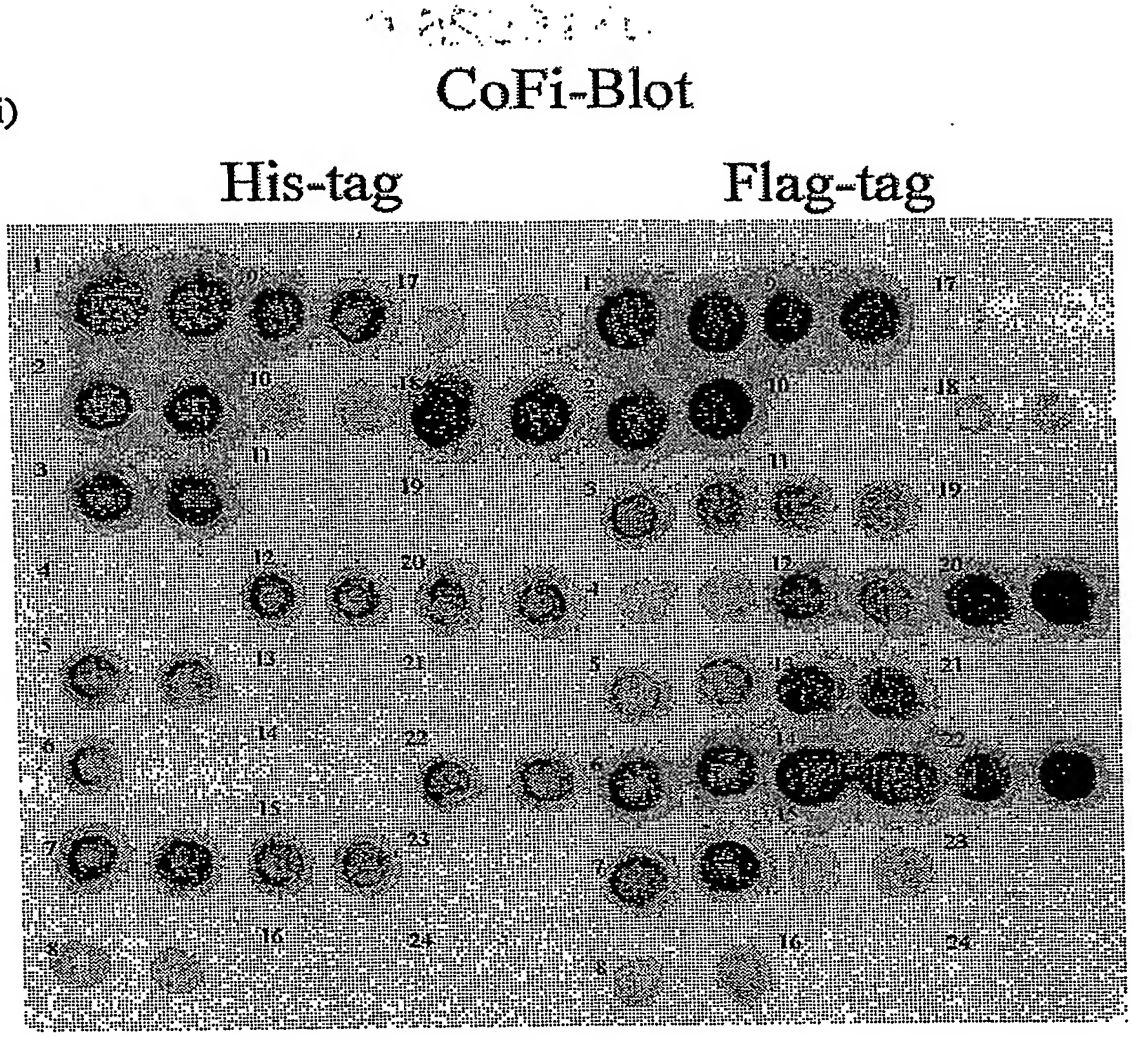


Figure 4b

(i)



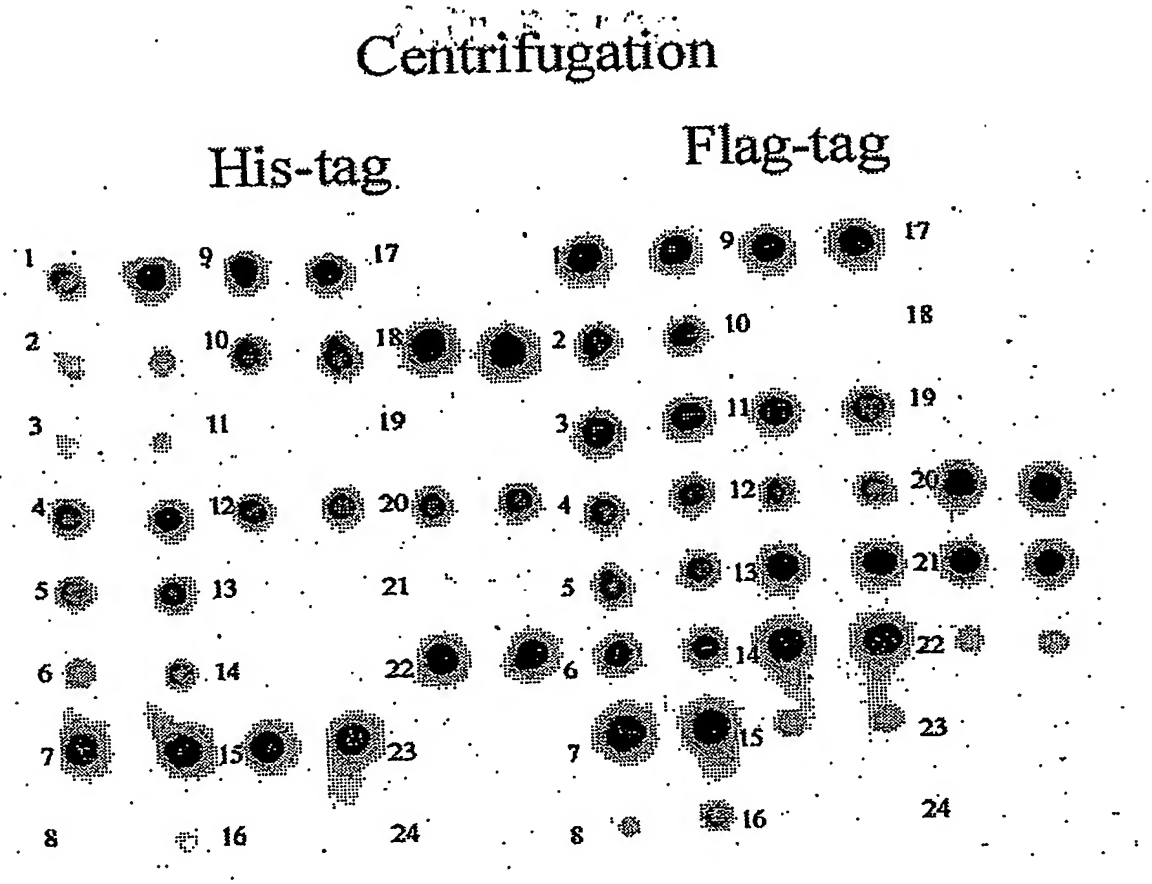
(ii)

	CoFi-Blot					
	His			Flag		
A	High	High		High	High	No
B	High	Low	High	High	No	No
C	High	No	No	Medium	Medium	No
D	No	Medium	Medium	Low	Medium	High
E	Medium	No	No	Medium	High	No
F	Low	No	Medium	High	High	High
G	High	Medium	No	High	Low	No
H		No	No	Low	No	No

False Positive or Negative
Predicted as soluble but in the most distant category
Predicted as soluble but in a neighbouring category
Correctly predicted in the right category

Figure 4c

(i)



(ii)

	Centrifugation					
	His			Flag		
A	High	High	No	High	High	No
B	Low	Medium	High	Medium	No	No
C	Low	No	No	Medium	Medium	No
D	Medium	Medium	Medium	Medium	Low	High
E	Medium	No	No	Medium	High	High
F	Low	No	High	Medium	High	Low
G	High	High	No	High	Low	No
H	No	No	No	Low	No	No

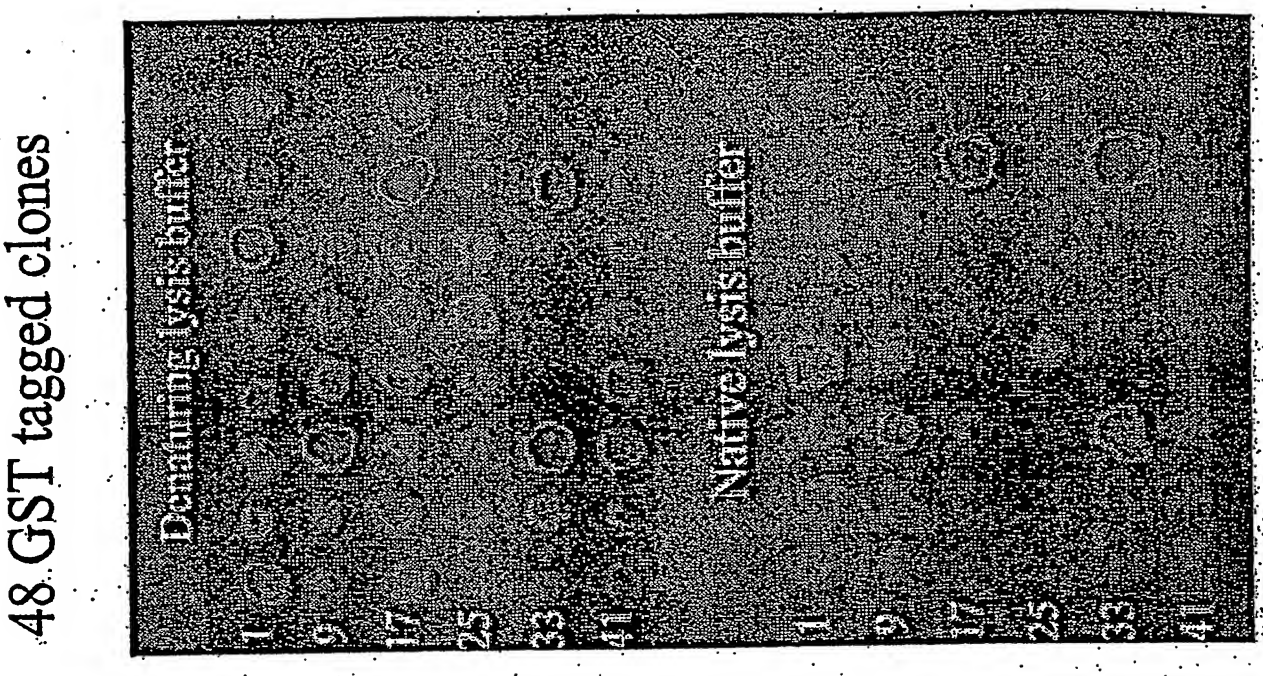

48 GST tagged clones
Denaturing lysis buffer
1
9
17
25
33
41
Native lysis buffer
1
9
17
25
33
41

FIGURE 5

Figure 6

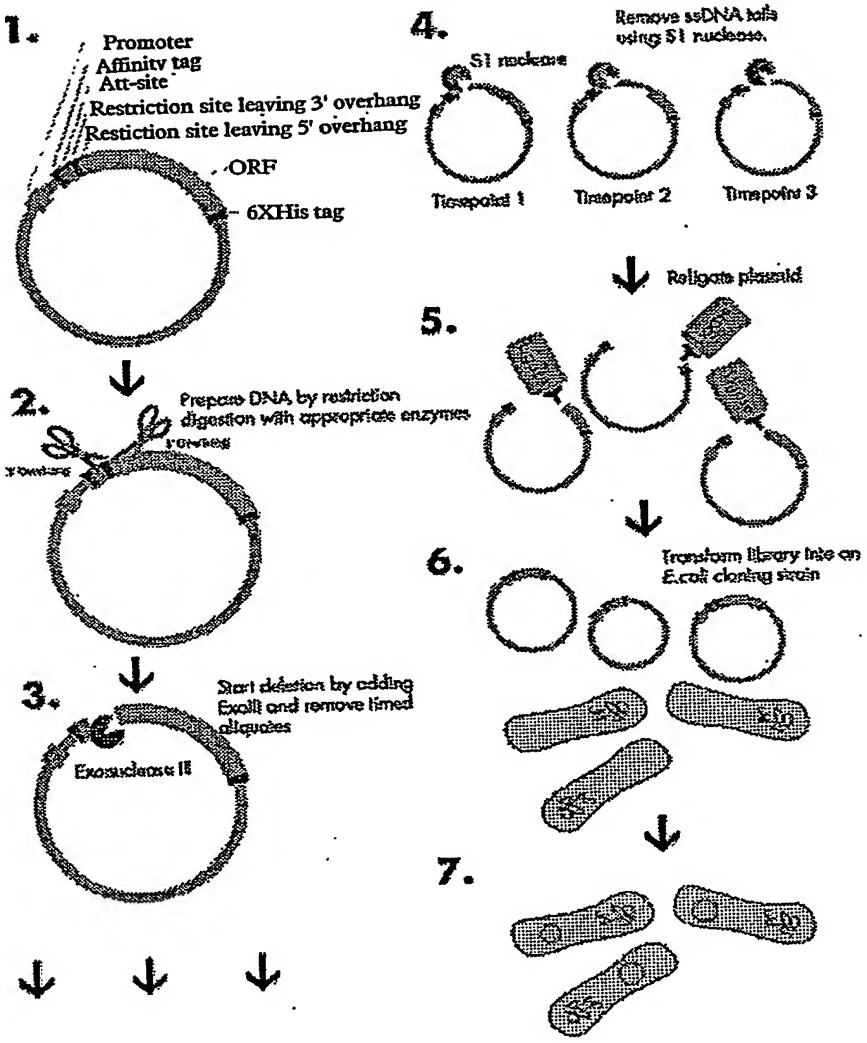
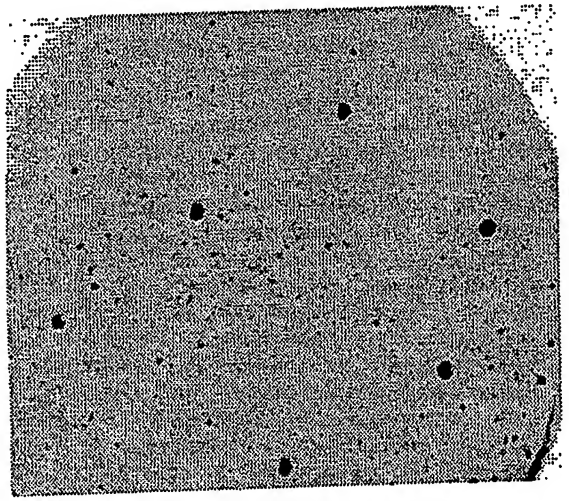
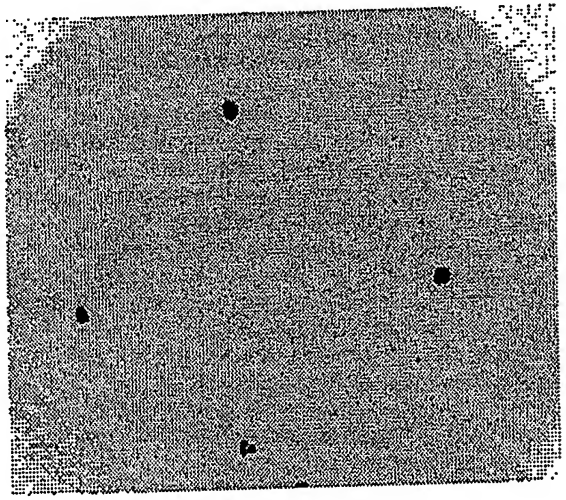

1.
Promoter
Affinity tag
Att-site
Restriction site leaving 3' overhang
Restiction site leaving 5' overhang
ORF
6XHis tag
2.
Prepare DNA by restriction digestion with appropriate enzymes
3.
Start deletion by adding Exo III and remove timed aliquotes
Exonuclease III
4.
Remove ssDNA tails using S1 nuclease.
S1 nuclease
Timepoint 1
Timepoint 2
Timepoint 3
Religate plasmid
5.
6.
Transform library into an E.coli cloning strain
7.

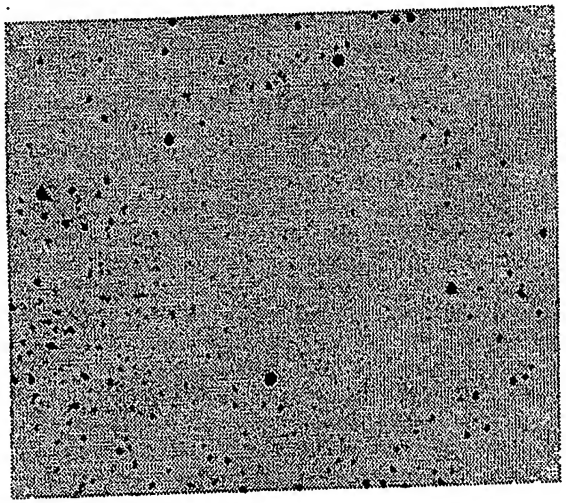
Figure 7a



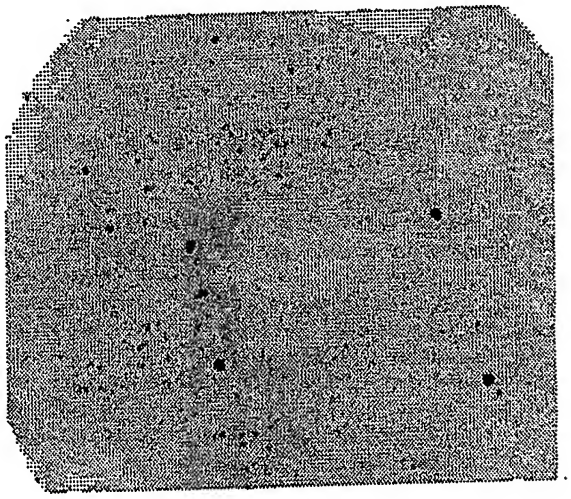
RORA 5-8 minutes
(~100-190 aa:s)



RORA 20-24 minutes
(~450-560 aa:s)

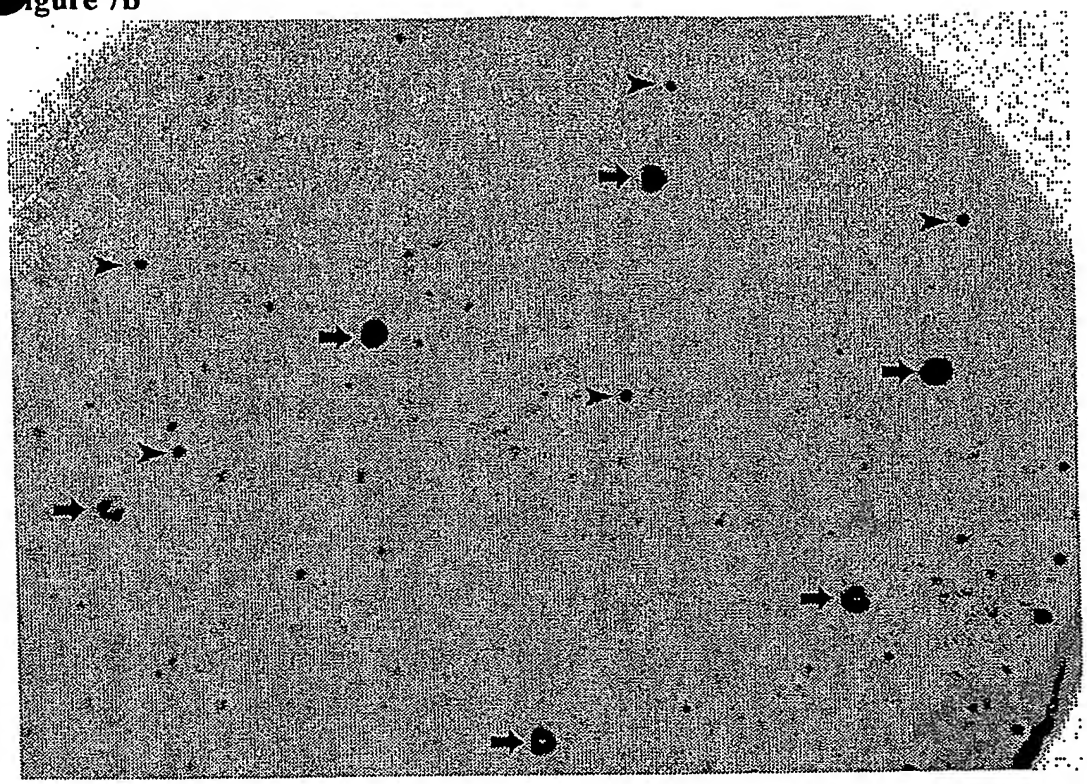


SOCS-2 1-4 minutes
(~0-90 aa:s)



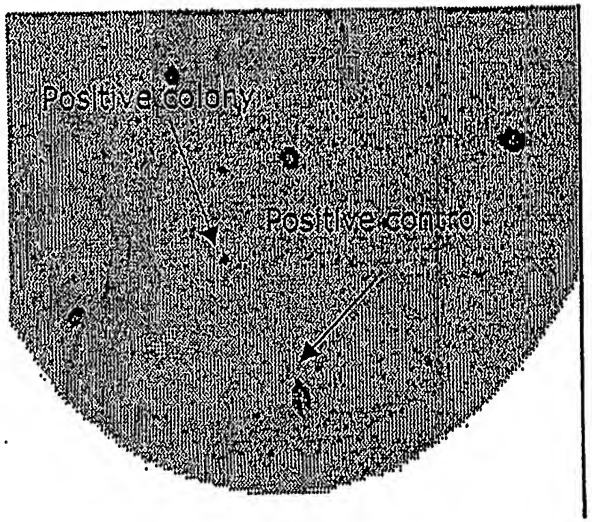
SOCS-2 9-12 minutes
(~210-300 aa:s)

Figure 7b

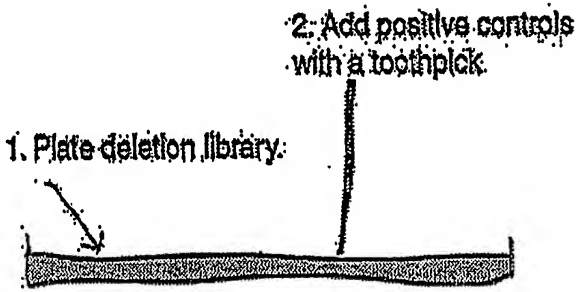


Colony screen of a RORa library. The positive controls (➡) are added to the plate with the library to facilitate orientation. ➢ Indicates examples of colonies judged to produce soluble protein.

Colony FiDo with reference



5 : 2500 Colonies



11/14

FIGURE 8

Figure 9

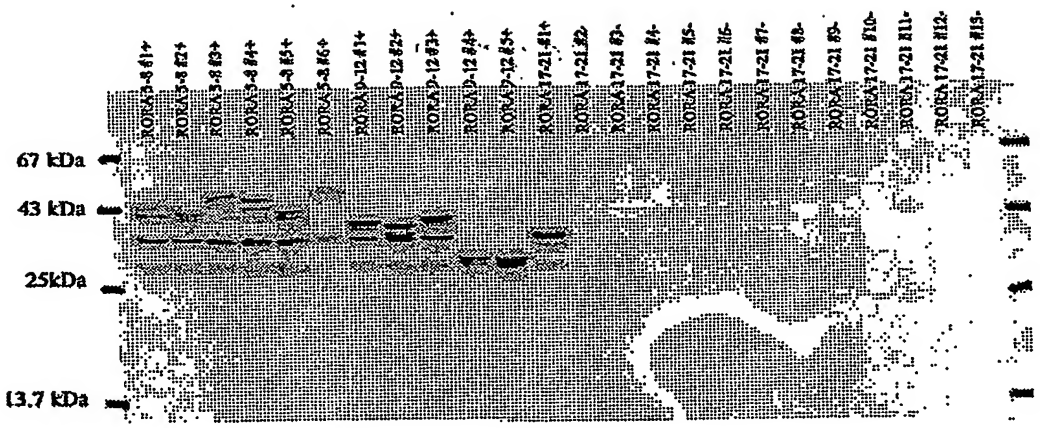


Figure 10

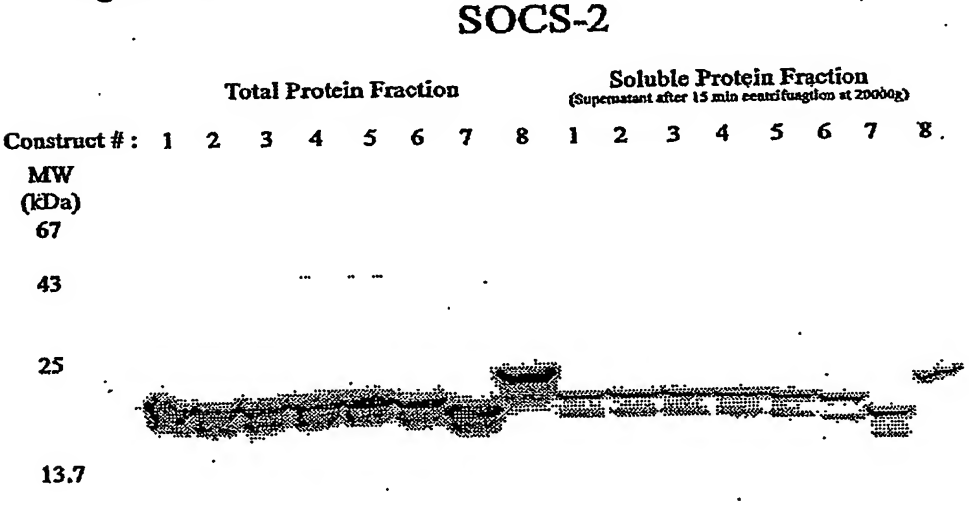
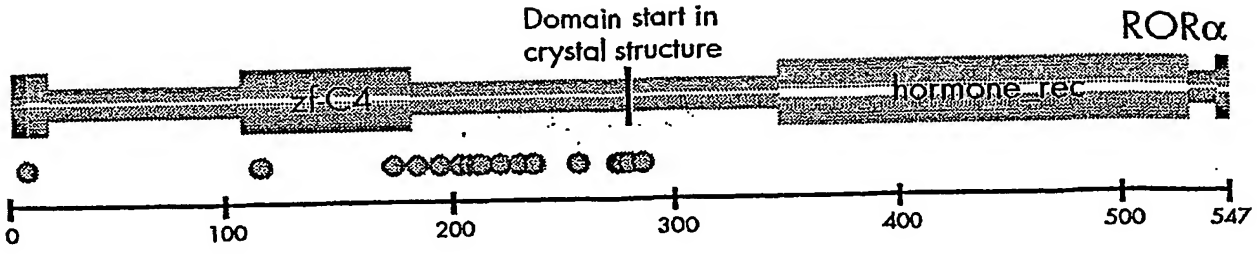


Figure 11



= Construct positive in the screen

= 6xHis-tag

= Flag-tag

= Att-site

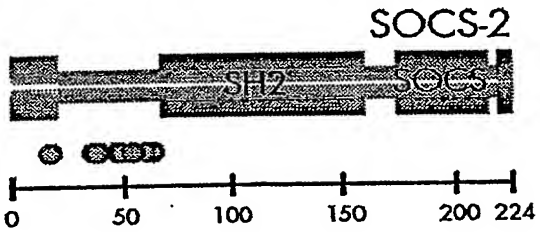


Figure 12